AD_____

Award Number: W81XWH-04-1-0404

TITLE: Large Scale Single Nucleotide Polymorphism Study of PD Susceptibility

PRINCIPAL INVESTIGATOR: J. W. Langston, M.D.

CONTRACTING ORGANIZATION: Parkinson's Institute
Sunnyvale, California 94089-1605

REPORT DATE: March 2005

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 2005	Annual (1 Mar 2004 - 28 Feb 2005)

4. TITLE AND SUBTITLE
Large Scale Single Nucleotide Polymorphism Study of PD Susceptibility

5. FUNDING NUMBERS
W81XWH-04-1-0404

6. AUTHOR(S)
J. W. Langston, M.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Parkinson's Institute
Sunnyvale, California 94089-1605

E-Mail: jwlangston@thepi.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

Although genetic contributions of Parkinson's disease (PD) have gained support from the recent identification of eight genetic loci in the familial PD, the results of intensive investigations of polymorphisms in dozens of genes related to sporadic, late onset, typical PD have not shown consistent results. Recent rapid progress in the investigation of single nucleotide polymorphisms (SNPs) has provided a new tool for this area of research. Millions of SNPs have been identified and compiled in several public accessible databases. A highly multiplexed genotyping technology called Molecular Inversion Probe Assay has recently been developed. This technique is capable of genotyping over 2000 SNPs in a single tube and is currently the most inexpensive platform for genotyping of SNPs with high call rate and high accuracy. Another major finding in recent genomic studies is that haplotype linkage disequilibrium is composed of blocks of sequence with average size of 7.8 kb and could be used in association studies. In this proposal, we plan to perform a large-scale association study by using the high throughput Molecular Inversion Probe Assay in PD to (1) investigate the association between classical, sporadic PD and 2386 SNPs in 23 genes implicated in the pathogenesis of PD; (2) construct haplotypes based on the SNP genotyping results to identify haplotypes associated with PD. This proposal is the first large-scale SNP association study in PD. The results of this study may lead to the further understanding of the pathogenesis of the disease and possibly new therapeutic approaches. In addition, experiences derived from this study may be applied in other complex disorders for the identification of susceptibility genes, as well as in genome-wide SNP association studies.

14. SUBJECT TERMS
Parkinson's disease genetics, SNPs, molecular inversion probe assay, haplotype analysis, genotyping

15. NUMBER OF PAGES
9

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION

There is an increasing consensus in the scientific community that Parkinson's disease is likely to result from a combination of environmental and genetic factors. In regard to the latter, a great deal of research has focused on finding functionally significant polymorphisms that predispose to the development of Parkinson's disease.

Several research groups have sought to identify PD susceptibility genes by carrying out population-based case-control association studies of candidate genes, identified on the basis of their involvement in a biologically relevant pathway that might be expected to lead to the degenerative process that underlies PD. However, no single genetic susceptibility factor has been shown to consistently increase the risk for PD (Mellick et al 1999, Nicholl et al 1999, Grevle et al 2000, Xu *et al* 2002, Zheng et al 2003, Le et al 2003, Marx et al 2003). These inconsistencies have prompted increasing concern about how future association studies should be performed (Nicholl et al. 1999, Markopoulou et al 1999). One problem is that most studies have investigated only one or a few polymorphisms in a particular candidate gene, with the assumption that the polymorphism would have an impact on the function/expression of the gene, at either transcriptional, translational or post-translational level, or that the polymorphism is in linkage disequilibrium with another nearby critical variant.

Unlike other studies, our proposal is designed to identify genetic susceptibility factor(s) for PD in a large group of PD cases and controls by genotyping 2386 SNPs in 23 related genes utilizing the Molecular Inversion Probe Assay - and to construct haplotypes based on the SNP genotyping results which will allow us to identify haplotypes associated with PD. Thus, we are far less dependent on correctly picking the right polymorphism from the study outset.

BODY

IRB: An IRB approval has been obtained for the proposal (please see attached document). Only non-Hispanic whites are included in the study solely for the purpose of reducing the genetic variation in the study subjects, which is critical to the study design. Similar studies on other ethnic groups will be conducted once this first study has been completed.

Sample collections: 200 PD patients and 200 controls are to be included in this study. The criteria for the patient inclusion are: (1) at least three of the four following cardinal features of PD on neurological examination: resting tremor, rigidity, bradykinesia, and postural instability; (2) a clear-cut response to levodopa; (3) an absence of clinical features suggesting atypical parkinsonism; (4) age of onset over 50 years; (5) a disease duration of 10 years or less (this criteria has been highly recommended by our epidemiologists to avoid survivor bias); (6) non-Hispanic white (to reduce the genetic complexity); (7) no familial history. The criteria for control subjects are non-Hispanic white without a history of known neurodegenerative disorders, matched in age, gender with the patient group. We have chosen to use primarily spousal controls. In our original proposal, we had planned to use samples that had been already collected. However, because of the manner in which these stored samples

had been collected, and the variable quality of the corresponding database, we have found it necessary to collect an entirely new set of samples so that we could (i) have absolute knowledge regarding clinical characteristics of the enrolled subjects and (ii) ensure the quality of the DNA. Taking this approach, we now have we now have collected 143 PD patient samples and 35 control samples that unequivocally meet the above described criteria. DNA quality has been checked for each with a spectrophotometer, and we have detailed information on ethnic origin, disease history, clinical presentation, drug response and family history in each case. Confidentially has been strictly maintained, as we are acutely aware of the serious hazards of the leaking of genetic information of study subjects and their families.

DNA purification: Genomic DNA from patient and control samples is extracted and quantified by spectrophotometer; DNA quality is verified by both Gel electrophoresis and spectrophotometer to check for DNA integrity and any possible contamination of DNA samples by RNA or/and protein. We require for all extracted DNA a 260/280 ratio between 1.7 to 2 and no degradation by electrophoresis gel, confirming the high quality of DNA.

SNP selection: The total number of SNPs found in several SNP databases (www.ncbi.nlm.nih.gov/SNP/, www.ensembl.org and www.snp.cshl.org) at the time of our proposal submission for each PD candidate gene is shown in table 1. In that table PD candidate genes were categorized as follows into three groups based on the number of SNPs found in each candidate genes. Group 1 contained candidate genes with less than 50 known SNPs; five genes (*GSTP1, CYP1A1, DRD1, DRD4* and *UCHL1* genes) were categorized in this group. Group 2 contained candidate genes with 50 to 300 SNPs, and contained 17. Group 3 included candidate genes with more than 300 SNPs; only one gene was categorized in that group. For the genes in Group 1 (containing less than 50 SNPs), we proposed to select DNA samples from a subset of patients for sequencing to find more SNPs for analysis. As part of this approach, we also periodically checked the SNP number within the 23 selected genes in the various databases (www.ncbi.nlm.nih.gov/SNP/, www.ensembl.org and www.snp.cshl.org) to include more polymorphisms in this study. Since that time many SNPs across the genome have been identified. The total number of SNPs currently identified (as 3/25/2005) for each PD candidate gene is listed in Table-1. As the number of identified SNPs in PD candidate genes has increased substantiall, we re-categorized PD candidate genes based on the SNP number found in each gene using the same definition that we used in our proposal (please see Table 1). As shown in Table 1, due to the accumulating number of SNPs identified, there are no longer any genes categorized in Group-1. Thus it will no longer be necessary to carry out the sequencing on a subset of Group 1 patients.

A critical component of this project is the selection of SNPs to be used, which has now been completed. We first selected SNPs for analysis from within regulatory regions, exons and splicing sites because they are likely to be more physiologically relevant. We then selected other SNPs such that the distance between 2 consecutive SNPs remained less than 1 KB. Once this was complete, the selected sequences for each gene were transferred to the system for probe tracking and designing.

Probe Tracking: We have transferred the 300 bases upstream and down stream of the SNP sequence to the system of oligo and probe tracking using a MS SQL server database with Java-based applications. This will serve to barcode and track all plates involved in the probe construction process as well as genotyping plates.

Probe Design: The ParAllele Probe Design Tool was used to design the probes. This tool includes repeat masker, genome blasts, melting temperature estimation, secondary structure analysis and other criteria to optimize probe design. Each molecular inversion probe is designed in two parts, which are put together after oligo synthesis to make the full length probe.

KEY RESEARCH ACCOMPLISHMENTS

- IRB approval was obtained for the proposal
- To date 143 PD patient and 35 control samples have been collected de novo that meet all inclusion criteria.
- Genomic DNA from all of these patients and control samples have been extracted.
- DNA for all collected samples was quantified by spectrophotometer and DNA quality was verified by both Gel electrophoresis and Spectrophotometer.
- We have and continue to search several SNP databases for all the SNPs in the above 23 genes selected for this study (www.ncbi.nlm.nih.gov/SNP/, www.ensembl.org and www.snp.cshl.org). The numbers of SNPs found in each gene are shown in table 2.

- Probe Design Tool was used to design the probes and probe Tracking file was made using a MS SQL server database with Java-based applications.

REPORTABLE OUTCOMES

While many of the milestones of this study have been met, we are still in the process of completing sample collection. Until this has been completed and the assays carried out will not have actual reportable outcomes.

CONCLUSIONS

This study is well under way and we anticipate having full enrollment within the next 3 to 6 months, and final results (including all newly added SNPs) within the next budget period. At that time we should have the complete data set analyzed and it will be possible to draw all relevant scientific conclusions.

REFERENCES

Grevle L, Guzey C, Hadidi H, Brennersted R, Idle JR, Aasly J. Allelic association between the DRD2 TaqI A polymorphism and Parkinson's disease. *Mov Disord* 15: 1070–1074, 2000.

Le WD, Xu P, Jankovic J, Jiang H, Appel SH, Smith RG et al. Mutations in NR4A2 associated with familial Parkinson disease. *Nat Genet* 33: 85–89, 2003.

Markopoulou K Langston JW. Candidate genes and Parkinson's disease. Neurology 1999;53:1382-1383

Marx FP, Holzmann C, Strauss KM, Li L, Eberhardt O, Gerhardt E, Cookson MR et al. Identification and functional characterization of a novel R621C mutation in the synphilin-1 gene in Parkinson's disease. *Hum. Molec. Genet.* 12: 1223-1231, 2003.

Mellick GD, Buchanan DD, McCann SJ, James KM, Johnson AG, Davis DR et al. Variations in the monoamine oxidase B (MAOB) gene are associated with Parkinson's disease. *Mov Disord* 14: 219–224, 1999.

Nicholl DJ, Bennett P, Hiller L, et al. A study of five candidate genes in Parkinson's disease and related neurodegenerative disorders. European Study Group on Atypical Parkinsonism. Neurology 1999;53:1415-1421.

Xu PY, Liang R, Jankovic J, Hunter C, Zeng YX, Ashizawa T et al. Association of homozygous 7048G7049 variant in the intron six of Nurr1 gene with Parkinson's disease. *Neurology* 58: 881–884, 2002.

Zheng K, Heydari B, Simon DK. A common NURR1 polymorphism associated with Parkinson disease and diffuse Lewy body disease. *Arch Neurol* 60: 722–725, 2003.

APPENDICES

No.	Gene	Number of SNPs at submission date (SNP#[1])	Projected SNP#[2]	Crrent number SNPs 03/23/05
1	GSTM1	56	56	58
2	GSTT1	55	55	72
3	GSTP1	18	50	62
4	GSTZ1	51	51	98
5	CYP2D6	52	52	69
6	CYP2E1	61	61	114
7	CYP1A1	29	50	58
8	NAT2	53	53	166
9	PON1	68	68	236
10	MAOA	54	54	134
11	MAOB	155	155	214
12	DAT	288	288	427
13	DRD1	18	50	69
14	DRD2	69	69	208
15	DRD3	62	62	157
16	DRD4	14	50	93
17	DRD5	51	51	114
18	COMT	153	153	238
19	VMAT2	164	164	223
20	NACP	201	201	426
21	Parkin	607	300	5282
22	UCHL1	11	50	68
23	tau	243	243	679

TABLE 1. (Group-1) genes with less than 50 SNPs, (Group-2) genes with 50-300 SNPs, and (Group-3) genes with more than 300 SNPs. PD candidate genes are categorized in table 1 as follows:

Group 1 in Yellow, Group 2 in Blue, and Group 3 in pink.

--Current SNP#1 indicates numbers of SNPs found in NCBI and the SNP Consortium websites at the time of proposal submission.

--Projected SNP#2 indicates current numbers of SNPs + numbers of SNPs we are planning to identify.

El Camino Hospital Medical Staff

2500 Grant Road, Mountain View, California 94040 (650) 940-7040 Facsimile (650) 966-9263

February 23, 2005

William Langston, M.D.
The Parkinson's Institute
1170 Morse Avenue
Sunnyvale, CA 94089

RE: Large Scale Single Nucleotide Polymorphism Study of PD Susceptibility (ECH-04-04)

Dear Dr. Langston:

The Institutional Review Board conducted its periodic review of the above-captioned protocol at its meeting of February 18, 2005 and recommended approval of continuation of the study. Per your report, there were no unanticipated side effects or research-related complaints. In addition, you continue to be in compliance with FDA regulations. The Board also approved modification to the protocol (2/7/05) and consent form (2/9/05).

Study Expiration Date: February 18, 2006

We look forward to another review in twelve months, or at the conclusion of your study, whichever occurs first.

Sincerely,

Michael S. Greenfield, M.D.
Chairman, Institutional Review Board
MSG/ae